ED 021 818 24 SP 001 626

By: Sorenson, Garth
A NEW ROLE IN EDUCATION: THE EVALUATOR.
California Univ., Los Angeles. Center for the Study of Evaluation of Instructional Programs.
Sponsoring Agency: Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No: BR-6-1646
Pub Date: Jan 68
Note: 4p.
Journal Cit: Evaluation Comment; v1 n1 Jan 1968
EDRS Price MF: $0.25 HC: $0.24
Descriptors: BEHAVIORISM, OBJECTIVES, CURRICULUM EVALUATION, EDUCATIONAL IMPROVEMENT, JOB ANALYSIS, *EDUCATIONAL OBJECTIVES, *EDUCATIONAL RESEARCHERS, EVALUATION CRITERIA, *PROGRAM EVALUATION

With the increase of federal funds for education, a new professional is emerging--the "evaluator"--whose role needs to be more clearly defined. He must not, first of all, take an absolutist position in his relationship to other educational experts or he will fail to get the cooperation of teachers and the support of powerful [illegible]. Second, he must accept certain basic assumptions with regard to educational goals: (1) educational goals should be defined in a process of interaction between professionals and representatives of the society, (2) goals and practices must be varied to accommodate a diverse population and must change as needs and values change, (3) goals must not be limited to purely academic objectives, and they must be stated in descriptive rather than interpretive language. The function of the professional evaluator should be to help teachers and administrators in a given school (1) to define their goals in terms of pupil performance, (2) to learn how systematically to discover differences among pupils that require particular kinds of instruction, and (3) to design and administer evaluation programs in order to find out which of their instructional procedures are paying off and which are not. It is hoped that the research and development evaluator will bridge the gap between the laboratory and the field by making explicit to the individual teacher the relevance of research findings to his work. (JS)

January 1968 Vol. 1, No. 1

BR-6-1646
PA-24

UCLA CSEIP

Evaluation comment

Center for the Study of Evaluation of Instructional Programs

a new role in education: the evaluator

With the increase of federal funds for education, a new professional is emerging – the evaluator. He is somewhat different from the expert in tests and measurements and in research design usually found working on a college faculty. Rather, he is a person who spends part or all of his working hours at research and development activities, thinking about and planning the evaluation of educational processes. Because his role is a new one on the educational scene, his functions and his relationship to other educational experts need to be more clearly defined. It is the aim of this article to present some ideas about that role.

Two papers on evaluation, one by Scriven (1965) and one by Stake (1966), contain a number of assertions and implicit assumptions about the evaluator's role which deserve examination. Among them are the following:

1. Scriven would assign evaluators the task of determining the effectiveness of instructional programs. But more than that, he would have them evaluate the goals of these programs as well. It is not enough for the evaluator to find out whether the teacher of mathematics or English or physical education has taught the students what he intended to teach them. The evaluator must also decide, Scriven believes, whether the specific course content was appropriate and worthwhile; for, as Scriven sees it, the evaluator is the person best qualified to judge.

2. Scriven holds that the relative goodness of different educational goals is to be determined by applying a set of absolute standards which will somehow be obvious to the evaluator. Apparently, Scriven doubts that it is possible for intelligent, informed, and well-intentioned people seriously to disagree about what should be taught, for he asserts that arguments over criteria turn out to be mainly "disputes about what is to be counted as good, rather than arguments about the straightforward 'facts of the situation,' i.e., about what is in fact good." (page 13)

3. Continuing his argument, Scriven implies that without absolute standards, evaluation is in fact probably impossible. "The process of relativism has not only led to over-tolerance for over-restrictive goals, it has led to incompetent evaluation of the extent to which these have been achieved . . ." (page 18)

4. Stake seems to imply that since absolute standards exist, it is not necessary to take the individual teacher's nor the individual school's goals into account. He seems to believe that such standards should be applied even if they relate only slightly or not at all to the local school's resources and goals. "It should be noted that it is not the educator's privilege to rule out the study of a variable by saying, 'That is not one of our objectives.' " (page 4, 11)

5. Both Scriven and Stake believe that it is possible and perhaps desirable to appraise teaching and other instructional programs independent of their effects on the stu-

continued

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED021818

SP001626

continued

dents. Stake (page 11) says, "The educational evaluator should not list goals only in terms of anticipated student behavior. To evaluate an educational program, emphasis must be given to what teaching as well as what learning is intended . . ."; and, "It is not wrong to teach a willing educator about behavioral objectives – they may facilitate his work. It is wrong to insist on them . . ." (page 12). Scriven further comments that ". . . pressure on a writer (curriculum maker) to formulate his goals, to keep to them, and to express them in testable terms, may enormously alter his product in ways that are certainly not always desirable." (page 21)

6. It may be inferred that Scriven believes that teachers who feel threatened by evaluators holding such absolute values should be ignored or at least discounted. "A little toughening of the moral fibre is required if we are not to shirk the social responsibilities of the educational branch of our culture." (page 5)

7. While it appears that he endorses most of Scriven's assertions, Stake would qualify at least one of them. If an individual evaluator were less than fully qualified, Stake would substitute a team of specialists as the appropriate determiners of educational goals and practices. The team would consist of experts in "instructional technology . . . psychometric testing and scaling . . . research design and analysis . . . the dissemination of information . . . (and perhaps) a social anthropologist" (page 23). He does not include historians, philosophers, businessmen, labor leaders, legal experts, or even non-behavioral scientists.

Te be sure, the assertions listed above do not constitute a summary of what Scriven and Stake have said in their papers. Nevertheless, it appears that they represent, at least roughly, some of the beliefs of Scriven and Stake and a point of view resembling that of a number of writers on public education.

"If evaluators . . . were to take an absolutistic position, a number of unfortunate consequences would follow"

In spite of the fact that a number of brilliant and famous men support a position similar to that just described, I believe that if evaluators generally were to take an absolutist position, a number of unfortunate consequences would follow.

For one thing, teachers would be unwilling to cooperate and work with these evaluators. An evaluator who insists on evaluating in terms of his own goals while ignoring what the school people are trying to do, an evaluator who criticizes them and the school for failing to do what they had not intended to do in the first place would certainly be viewed as threatening. It can be safely predicted that teachers who feel threatened will resist and will devote their time and energies to defending old practices rather than to examining and improving them.

A second unfortunate consequence would be that evaluators would not get the support they need from powerful groups in the community who have a legitimate interest in what goes on in the school. Evaluation requires large amounts of time, money, and other commodities that evaluators cannot get without a good deal of public support – especially if they already have alienated the teachers and school administrators. Many of the individuals and groups in this country whose support is needed believe that the schools were invented to serve the needs of society and ultimately are answerable to the taxpayers, or at least to someone other than professional evaluators.

These individuals and groups do not always agree with one another about how the schools can best serve society, but they do agree that the schools are not autonomous. Many of these individuals – for example, Paul Goodman, Robert Hutchins, Sidney Hook, James Conant, John Goodlad, Roald Campbell, Ralph Tyler, Clark Kerr, Admiral Rickover, Harold Taylor, Paul Woodring, Jerome Bruner, David Ausubel, Myron Lieberman, Lawrence Cremin, Benjamin Bloom, to name only a few, as well as many groups – have given a good deal of thought and study to questions about the goals and methods of education. They are likely to regard individuals whose main qualification for the prescribing of educational goals is that they are experts in psychometry, research design, or social anthropology, but who are ignorant of the philosophical and political issues in education, as naive, arrogant, parochial, and, therefore, unworthy of assistance.

A third possible consequence – an evaluation program based upon the absolutistic assumption that "good" educational programs exist independent of persons and their preferences and independent of what students learn – is bound to fail. Its results are certain to be inconclusive and meaningless.

An analogy can be found in the attempts to evaluate teacher effectiveness. After surveying the results of half a century of research, investigators like Anderson and Hunka (1963) and Turner and Fattu (1960) have concluded that research in this area has been unproductive and has reached a dead end because of problems encountered in developing suitable criterion variables. In statistical terms, the variables lack reliability. It is my contention that the reason for the failure to develop usable criterion variables is a basic error in the way in which the researchers conceptualized the problem – more specifically, in their reliance on an absolute model of teacher effectiveness. Virtually all the investigators assumed either implicitly or explicitly the existence of sets of behaviors that objectively define the teacher – behaviors which exist as an absolute, independent of any particular observer and which would be recognized by an experienced educator when he encountered them, even though he might not be able to verbalize them in advance. Those researchers were failing to recognize and take into account the fact that any two observers are likely to differ in their beliefs about the ideal traits of the good teacher.

Ryans (1960) found that even when two observers were simultaneously watching the same teacher, they did not agree about him in their independent ratings unless they had had considerable training in Ryans' rating system – and sometimes not even then. It was probably his observers' differing notions about the ideal teacher they were observing. Analogously, any two evaluators are likely to disagree about the goals of education and can, therefore, be expected to disagree about the "goodness" of whatever actual method or program they may at a specific time be seeking to evaluate. The point is, there never has been and never will be general agreement on the goals of education any more than there is agreement on the qualifications and characteristics of the ideal teacher. Though particular groups of people will agree on particular goals, we must

live with the fact that there is a welter of conflicting ideas on the subject in the society as a whole.

"a set of assumptions which may provide a reasonable alternative"

Following is a set of assumptions which may provide a reasonable alternative to those selected from Scriven and Stake.

1. Educational institutions should serve the needs of society and of the individuals who comprise it; these needs are complementary and interdependent.

2. A society's needs can best be defined by the members of that society through discussion, persuasion, and, ultimately, through voting. To insure that the goals of education will correspond with the citizens' views of their needs, the goals should be defined in a process of interaction between professionals and representatives of the society.

3. Every society changes; its needs and values are in a constant state of flux. Because of increases in population, knowledge, and technology, our society is very different from what it was even a decade ago. We now need new classes of workers, e.g., technicians who can build and operate computers. And because, as Gerard Piel (1961) has pointed out, we are no longer a society characterized by scarcity of goods, values based on dearth, such as hard work, thrift, etc., are less salient. Concomitantly as our needs and values change, we must expect our educational goals to change.

4. Even though many of our values seem to be changing, we continue to prize diversity. Ours is a pluralistic society with different religions, political viewpoints, subcultures, and values. We believe that our heterogeneity makes our society richer, more interesting, and stimulating. What is even more critical, we believe that heterogeneity makes our society viable. To accommodate such a diverse population, we must expect our educational goals and practices to be varied.

5. The goals of our educational institutions are not and never have been limited to purely academic objectives. Most people want the schools to do more than to teach the traditional academic subjects: they want individual and societal objectives included. For example, a century ago, the McGuffey Readers attempted to inculcate moral principles. More recently, James B. Conant (1953, page 62) said that the schools should provide a basis for the growth of mutual understanding between the different cultural, religious, and occupational groups in our country. "If the battle of Waterloo was won on the playing fields of Eton, it may well be that the ideological struggle with Communism in the next fifty years will be be won on the playing fields of the public high schools of the United States."

6. We can tell if an educational program or teaching method is working only by observing whether hoped-for changes are occurring in the students – while at the same time making certain that damaging changes are *not* occurring, e.g., learning to hate a particular subject, or learning to believe one cannot learn arithmetic even if he works at it. We cannot properly evaluate an instructor or a program without assessing the effects, wanted and unwanted, on students. To evaluate a schedule of events within a school, or a series of teacher activities, or any array of teacher characteristics while neglecting the product is to examine intentions without considering consequences.

7. Educational goals must be stated in descriptive rather than in interpretive language. We have learned that it is not useful to define educational goals in the terms formerly used by professional educators and still used by their critics. We know that instead of such high-sounding slogans as "transmitting the cultural heritage," "educating citizens for democracy," and "developing the individual's potential," we must develop objectives defined in terms of changes in pupils' behavior or in the products of student behaviors. We must also be careful that, in rigorously setting behavioral goals, we do not slip into triviality. We must be prepared to defend each behavioral goal in terms of value assumptions and to answer the question why one particular behavioral goal is better than another. These points do not represent new thinking. They describe a trend, which according to Ralph Tyler (1954, 1956) began about 1935, a trend of which many public school teachers still are unaware. Tyler stated that it is more important to evaluate the educational process than the structure of the school and that it is more important to evaluate the product than the process. I would rephrase this point: the proper way to evaluate both the educational process and the structure of the schools is to find out whether they are in fact producing the hoped-for product.

"the function of the professional evaluator"

The function of the professional evaluator should be to help teachers and administrators in a given school to do such things as the following:

1. Define their goals in terms of pupil performance. John McNeil (1966), director of Supervised Teaching at UCLA, and I both have found that many experienced teachers are not able to define their objectives in language which describes observable changes in pupil behavior. It is easy to be critical of such teachers, and it is easy to state educational goals behaviorally – if we limit ourselves to rote learning. For example, "students will be able to name the bones of the body" is a goal stated in behavioral terms. While this goal may be important in some contexts, it is a very limited one. The behavioral definition of higher order goals is much more difficult. At the end of a course, teachers want their students to perform in such a manner as to warrant the inference that the students have learned to "know," "understand," "appreciate," and "think" about what the teacher has tried to teach. Merely to tell teachers that they should state these goals behaviorally is far from sufficient. What would be more helpful would be to show them how, and to invent more sophisticated instruments for them to use.

2. Learn how systematically to discover differences among pupils that require particular kinds of instruction. Teachers need appraisal devices that will do more than reveal differences in what students already have learned. They need instruments that will also reveal barriers to,

continued

continued

or interferences with, learning, among them, (a) misconceptions; (b) particular habits, such as failure to pay attention; (c) certain needs that the child is satisfying at the expense of learning, e.g., need for group approval or sensitivity to peer pressures; and (d) attitudes deriving from class and ethnic background, etc. Some important differences among students are so subtle that, without sophisticated instruments, the child who has not learned to attend to the teacher's instructions may be mistaken for a dull child, or an angry one, or perhaps one with a constitutional impairment.

3. Design and administer evaluation programs. More importantly, professional evaluators should help individual teachers to find out which of their instructional procedures are paying off and which are not. With guidance, it is possible for the teachers themselves to try out and to evaluate alternative instructional methods on the job. For example, Bartlett (1960) demonstrated that when an instructor spent part of his time in an algebra class teaching study habits, the students learned more than when he spent the entire time teaching algebra.

Scriven, M. *The methodology of evaluation.* Bloomington: Indiana University, 1965. (mimeographed paper)

Stake, R. E. *The countenance of educational evaluation.* Urbana: University of Illinois, Center for Instructional Research and Curriculum Evaluation, 1966. (mimeographed paper)

Tyler, R. W. "Modern aspects of evaluation." *California Journal of Secondary Education,* 1954, 29, 410–412.

Tyler, R. W. "The curriculum: then and now." *Proceedings, 1956. Invitational Conference on Testing Problems.* Princeton, New Jersey: Educational Testing Service, 1957.

"hopefully,
the research and development
evaluator will bridge the gap"

Public school people do not need more critics — critics abound. What these educators *do* need is someone to help them find and test alternative solutions to the complex problems they face daily. For the most part, university personnel who have the knowledge to perform the kinds of evaluation functions described above have not been taking their knowledge to the schools. They have been publishing their findings in professional journals, but they have failed to make explicit to teachers the relevance of those findings for the teachers' work. Hopefully, the research and development evaluator will bridge the gap between the laboratory and the field.

Garth Sorenson
Associate Professor of Education
University of California, Los Angeles

references

Anderson, C. C., and Hunka, S. M. "Teacher evaluation: some problems and a proposal." *Harvard Educational Review,* 1963, 33 (1), 74–96.

Barlett, W. H. "The practical application of psychological facts in the classroom as demonstrated by teaching a specific study method to elementary algebra students of the junior college level." Unpublished master's thesis, University of California, Los Angeles, 1960.

Conant, J. B. *Education and liberty.* Cambridge: Harvard University Press, 1953.

Piel, G. *Science in the cause of man.* New York: Random House, 1961.

Ryans, D. G. *Characteristics of teachers.* Washington, D. C.: American Council on Education, 1960.